Aplicación del algoritmo de Watershed para análisis de calidad de arenas para fracking

Carlos Gustavo Rodriguez Medina, Oscar Daniel Chuk, Adriana Luna, Regina Bertero, Enrique Núñez, Pablo Trigo

Instituto de Investigaciones Mineras, Facultad de Ingeniería – Universidad Nacional de San Juan (UNSJ)
Código Postal J5400ARL. Av. Lib. Gral San Martin 1109 – San Juan – Argentina

{grodriguez,dchuk}@unsj.edu.ar

Abstract. *The present work is part of the digital image processing area. The Watershed transform is applied in order to segment the elements of interest into specific digital images. Although the algorithm that implements this transformation is commonly used in the processing of biomedical images, for this case it is used in an application for industrial purposes, in order to analyze the quality of sands used in Fracking processes (fracturing hydraulic rock) for the extraction of hydrocarbons in an unconventional way.*

Resumen. *El presente trabajo se enmarca en el área de procesamiento digital de imágenes. Se aplica la transformada de Watershed (de su traducción en español conocido como de divisoria de aguas) con el fin de segmentar los elementos de interés en imágenes digitales específicas. Si bien el algoritmo que implementa esta transformada suele ser utilizado comúnmente en el procesamiento de imágenes biomédicas, para este caso se hace uso en una aplicación con fines industriales, con el objeto de poder analizar la calidad de arenas utilizadas en los procesos de Fracking (fracturación hidráulica de rocas) para la extracción de hidrocarburos de forma no convencional.*

1. Introducción

Watershed es una técnica de segmentación basada en morfología matemática, que permite extraer las fronteras de las regiones que hay en una imagen. A la vez, se considera una técnica de segmentación basada en regiones, debido a que clasifica los pixeles según su proximidad espacial, el gradiente de sus niveles de gris, y la homogeneidad de sus texturas. Por ello se toma como una técnica de detección de contornos y crecimiento de regiones al mismo tiempo [Gonzales y Ballerin 2008], [La Serna y Garcia 2011].

A través de Watershed se puede considerar una imagen en escala de gris como la imagen topográfica de un relieve terrestre; en donde a cada pixel se le asocia como valor de altura su nivel de gris correspondiente. En este sentido, se puede pensar que las intensidades de gris de mayor amplitud se corresponden con montañas, mientras que las intensidades de menor valor se corresponden con valles y ríos [La Serna y Garcia 2011].

La técnica además incorpora un proceso de inundación de los valles, desde los niveles más bajos de altura (valores mínimos locales, que constituyen cuencas de inundación rodeadas por cadenas montañosas), hasta los niveles más altos.

Las zonas de baja intensidad de gris también se conocen como vasijas (basins en Inglés), por donde fluirá el agua e inundará toda la topografía de la imagen; es decir, el agua fluirá en cada una de las vasijas identificadas. El proceso de inundación continúa hasta que las aguas de cuencas o vasijas contiguas se unen, formando líneas de unión que representarán las fronteras de regiones homogéneas, y constituyen el resultado de la segmentación. Lo descripto se presenta en la figura 1.

Las líneas divisorias de aguas forman los contornos de los objetos de la imagen que se quiere analizar, que darán lugar a la segmentación de la imagen, y que corresponden además a zonas de elevada intensidad de gris.

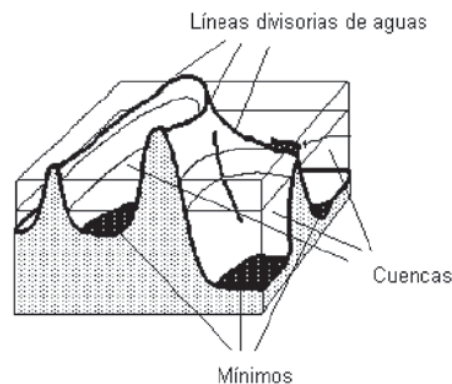


Figura 1. Watershed: Líneas divisorias, mínimos y vasijas o cuencas. Fuente [La Serna y Garcia 2011].

Se debe tener especial cuidado durante la aplicación del algoritmo de Watershed, dado que en imágenes no homogéneas y/o embebidas en ruido se genera un gran número de mínimos locales, produciéndose sobresegmentación en pequeñas regiones, en donde muchas de las regiones generadas no son importantes dentro de la imagen, o no representan objetos existentes en la imagen original. Esto genera errores en las mediciones o propiedades de los elementos de interés de la imagen.

El algoritmo de Watershed permite la segmentación de estructuras complejas que no pueden ser procesadas mediante otros métodos convencionales de procesamiento digital de imágenes. La transformada de Watershed es fácilmente adaptable a los diferentes tipos de imágenes y es capaz de distinguir objetos sumamente complejos que no pueden ser analizados correctamente mediante algoritmos convencionales [Gonzales y Ballerin 2008].

En este orden de ideas, también resulta ser de utilidad el algoritmo de Watershed en el análisis de elementos microscópicos, como lo es para el caso partículas o granos de arenas utilizadas para la fracturación hidráulica.

El proceso de Fracking o el Fracturamiento Hidráulico, es el proceso utilizado en la extracción de hidrocarburos de manera no convencional, como es el caso del yacimiento de Vaca Muerta, en la provincia de Neuquén. En dicho proceso se aplica presión hidráulica a la roca reservorio de hidrocarburo, hasta que se produce una falla o fractura de la misma mediante la inyección a alta presión de fluido viscoso, y se mantiene abierta con un agente apuntalante (tal como arena natural o sintética) una vez que se haya liberado la presión de inyección.

El agente de sostén, agente apuntalante o propante es un material usado para evitar que la fractura generada en la roca durante el Fracturamiento Hidráulico se cierre completamente, una vez terminado el bombeo y garantizar una buena conductividad al canal recién creado.

A fin de garantizar la capacidad de las arenas con destino al Fracking para actuar como apuntalantes o propantes, se reconocen como base las normas de referencia del American Petroleum Institute (API). Esta norma recomienda una serie de ensayos a realizar sobre las arenas con el fin de conocer la calidad del material a utilizarse en el Fracking. Entre Los ensayos a realizarse se encuentran los de Redondez y Esfericidad de los granos de arena.

Los ensayos de Esfericidad y Redondez padecen de una fuerte dependencia del criterio del observador, dado que el procedimiento se realiza mediante la utilización de un microscopio y las mediciones de esfericidad y redondez se realizan por comparación respecto a tablas que contienen una serie de formas provistas por la norma. Esto deriva en resultados sustancialmente disímiles que invalidan el procedimiento, aunque es el que se sigue usando por norma.

El uso de técnicas de Visión Artificial se presenta como la alternativa tecnológica que puede permitir obtener medidas objetivas no dependientes del criterio de un operador.

2. Desarrollo

Al operar con imágenes correspondientes a muestras de arenas, los elementos de interés se esparcen sobre una superficie contrastante (utilizada como fondo), teniendo la precaución de evitar el amontonamiento de las partículas, ya que cuando se captura la imagen y al quedar partículas unidas unas a otras un proceso de segmentación sencillo (como una umbralización) no es suficiente. Por mayor cuidado que se tenga, generalmente se produce algún efecto de amontonamiento o aglomeración de las partículas de arena, es prácticamente inevitable.

Por tal motivo, resulta de interés encontrar algún mecanismo para salvar dicha situación, dado que para estos casos la segmentación no arroja resultados satisfactorios, ya que se suelen obtener elementos que no representan las partículas de arenas reales.

En el marco de este trabajo, se han probado distintos algoritmos de segmentación, pero con el que mejores resultados se ha podido obtener ha sido mediante el algoritmo de Watershed.

Esto, de alguna manera confirma lo expresado por [Gonzales y Ballerin 2008], que indica que la partición de una imagen en regiones permite obtener información que busca el observador mediante una herramienta morfológica poderos denominada transformada Watershed. También manifiesta que "esta transformada es fácilmente adaptable a los diferentes tipos de imágenes y es capaz de distinguir objetos sumamente complejos que no pueden ser analizados correctamente mediante algoritmos convencionales" .

A continuación se expone la aplicación del algoritmo de Watershed utilizado para poder segmentar granos de arena en imágenes digitales, y así poder realizar posteriormente un análisis de calidad, a través de sus propiedades de Redondez y Esfericidad (dichas propiedades no son tratadas en este trabajo por la extensión del mismo).

Para el procesamiento digital de las imágenes, se utiliza MATLAB. Este es un software matemático que ofrece un entorno de desarrollo integrado (IDE) con un lenguaje de programación propio (lenguaje M). Está disponible para las plataformas Unix, Windows y Apple Mac OS X. [Matworks 2018].

Está orientado para llevar a cabo computaciones numéricas casi de todo tipo, pudiendo manipular vectores y matrices tanto reales como complejas con funciones y fórmulas de variadas ramas de la matemática. Matlab se compone de un programa básico y un conjunto de Toolbox para labores más especializadas [Matworks 2018].

La ventaja de operar imágenes con Matlab, es que las mismas se cargan al entorno de trabajo en forma de matrices, y dado que este software está orientado a aplicar toda su potencialidad de cálculo a las matrices, es que resulta una herramienta potente para la manipulación y procesamiento matemático; además de poder procesarse digitalmente las imágenes a través del Toolbox de Procesamiento de imágenes con el que cuenta (por ej. aplicar filtros, mascaras, transformadas, etc.).

Tal como se ha mencionado anteriormente, a través del Toolbox de procesamiento de imágenes, Matlab dispone de la posibilidad de aplicar la transformada de Watershed, implementada mediante la función "watershed()".

El algoritmo elaborado en este trabajo, a su vez utiliza dicha función o algoritmo. Para su utilización de modo apropiado, se realizan taras de pre-procesado de la imagen de interés. Tal como se presenta mediante la figura N° 2, el diagrama de flujo expone de forma resumida el algoritmo resultante.

En relación a la figura 2, como primer paso se carga en el espacio de trabajo de Matlab la imagen de interés a ser procesada. Se dispone de una función específica para ello, "imread()", la que permite que la imagen digital, almacenada en forma de un archivo en el disco de la PC, se cargue en memoria como una matriz, de NxMxR. Donde N hace referencia a la cantidad de filas, mientras que M a la cantidad de columnas y R al número de planos de color de la imagen, es decir a los planos de color R(rojo), G(verde) o B(azul) de una imagen de 24 bpp color [Rodriguez. y Navas 2014].

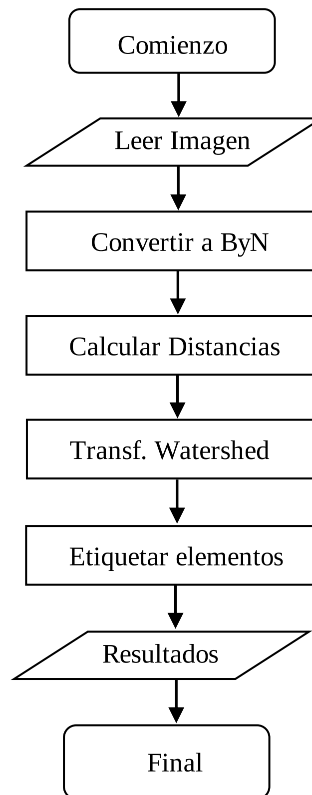


Figura 2. Algoritmo para procesar imágenes de arenas a través de Watershed.

Luego la imagen se convierte en binaria, o blanco y negro, para así poder aplicarle una transformada de distancia. Esta es a los fines de poder encontrar elementos que definirán las regiones o vasijas para la posterior aplicación del algoritmo de Watershed [Fiter 2012].

Cabe destacar que Matlab permite convertir imágenes de color a binario (blanco y negro) y a escala de grises de manera muy rápida y sencilla para el operador, mediante simples funciones específicas.

La transformada de distancia de una imagen binaria, es la distancia de cada pixel al pixel más cercano distinto de cero [Fiter 2012]. Suele utilizarse en operaciones morfológicas tales como la erosión, dilatación y en la segmentación.

También se utiliza en forma conjunta con la transformada de Watershed para realizar la segmentación, para dividir formas convexas que se tocan entre si y poder considerarlas como elementos separados.

Para el cálculo de distancia, se pueden utilizar distintas métricas, haciendo uso de métodos tales como: Euclídea, City block, Chessboard y Cuasi-euclídea.

Como ya se había anticipado, la transformada de Watershed es una técnica de segmentación que interpreta la imagen como un relieve topográfico, en el que la intensidad de cada pixel se considera como su altura en dicho relieve y calcula las líneas divisorias, las que representan los limites de las cuencas hidrográficas o valles del relieve.

Matlab implementa dicha transformada a través de la función de igual nombre, "watershed()". Esta produce como resultado una matriz con las mismas dimensiones que la matriz de entrada (sería la imagen que se está procesando), en la que se identifican las regiones obtenidas mediante Watershed.

Para el caso del presente trabajo, se utiliza el algoritmo de Watershed en conjunto con la transformada de distancia de los pixeles de la imagen. Se debe tener en cuenta que la utilización de estos algoritmos puede llegar a producir algún nivel de sobresegmentación en la imagen de resultado. Es decir, pueden aparecer una mayor cantidad de líneas divisorias que la cantidad de elementos a segmentar, subdividiéndose dichos elementos de interés, resultado esto en un efecto no deseado.

Dependiendo de los resultados obtenidos respecto a los deseados, se puede emplear el algoritmo de Watershed en conjunto con la transformada de distancia, o con otras técnicas.

También se debe considerar la necesidad de realizar algún filtrado previo (u otro pre-procesamiento) a la imagen, para eliminar ruido o impurezas, evitando así una posible sobresegmentación. Para este caso, referido al desarrollo del presente trabajo, se ha empleado la transformada de distancia en conjunto con Watershed, con un filtrado previo. Esto se ha utilizado así, al menos en una primeara instancia, conociendo que existen alternativas para evitar los inconvenientes de la sobresegmentación, en caso de ser necesario.

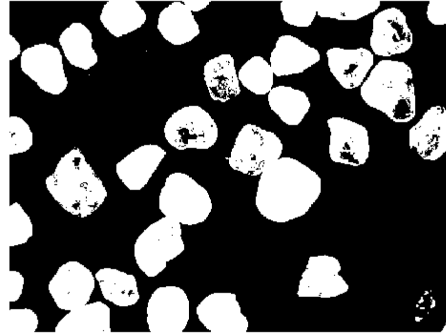
3. Resultados y conclusiones

Tal como se describió en los apartados anteriores, el propósito es el de poder procesar digitalmente imágenes de arenas realizando una segmentación adecuada, con el objetivo de contar con los elementos (correctamente aislados) de interés de las partículas de arena en estudio. Y así luego realizar diversas mediciones (esfericidad y redondez) referidas a determinar la calidad del material que será utilizado en procesos de fracturación hidráulica.

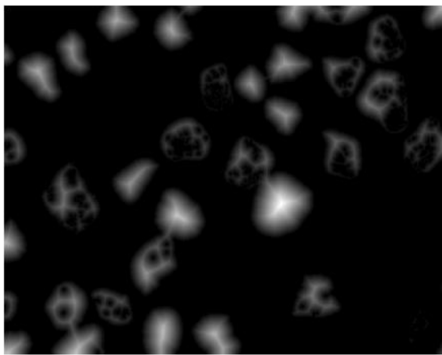
En la figura 3, se exponen una serie de imágenes, las que corresponden a un ejemplo concreto de uso del algoritmo elaborado. La primera de ellas, se refiere a una imagen original (sin ningún procesamiento) de una muestra de partículas de arenas, utilizada como entrada al algoritmos de la figura 2. La figura 3-B, corresponde a la imagen convertida a binario. Para el caso de la figura 3-C, se presenta el resultado de la transformada de distancia de la imagen binaria, para luego aplicarle la transformada de Watershed y así obtener la imagen expuesta en la figura 3-D.



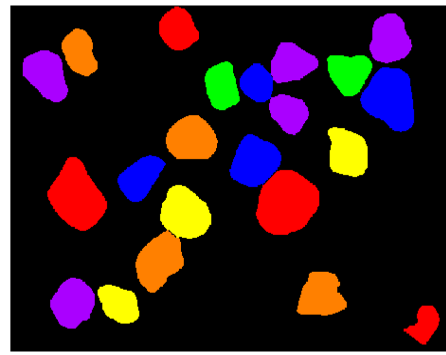
A - Imagen original de particulars de arena.



B - Imagen convertida a binario.



C- Imagen resultado de la transformada de distancia.



D - Imagen resultado de aplicar Watershed.

Figura 3. Resultados por etapas al aplicar el algoritmo de la figura 2.

La imagen correspondiente a la figura 3-D, es la de mayor interés como resultado obtenido. En ella se aprecian las partículas de arena segmentadas correctamente, aun para los casos en los que se hallan partículas que se tocan entre sí. Nótese que se hace referencia a una segmentación correcta o adecuada, desde el punto de vista que no se presentan casos en que el algoritmo confunda dos partículas pegadas como si fuera una sola, o también algún otro defecto, tal como el de sogresegmentación. Se aprecia la utilización de diferentes colores en las partículas de arena, esto es presentado así por el propio algoritmo elaborado, y es a los fines demostrativos de que el algoritmo las ha considerado como elementos diferentes, logrando así una segmentación adecuada. El algoritmo etiqueta cada uno de los elementos obtenidos (segmentados) para luego poder operar sobre ellos en forma independiente, y poder aplicar las métricas que correspondan.

Se debe tener en cuenta, que como parte del preprocesamiento, previo a realizar el cálculo de la trasformada de Watershed, a la imagen se le ha eliminado algunas impurezas (ruido), como así también aquellas partículas que estaban "pegadas" al contorno de la imagen, ya que si estas partículas se segmentan podrían tratarse de partes o fracciones solamente, introduciendo errores a las posteriores mediciones a realizar.

Se han llevado a cabo diversas pruebas del algoritmo elaborado, y para las imágenes de interés utilizadas se han podido obtener resultados satisfactorios. Pero se

sabe que la sobresegmentación es un problema latente, y que se deben buscar otras alternativas para tratar de eliminarla o minimizarla.

Entre varios algoritmos utilizados empleando otros tipos de segmentación, este ha sido el que mejores resultados a permitido obtener. También, como parte de tareas a seguir desarrollando, se pretende estudiar la utilización de la transformada de Watershed en combinación con otras técnicas para reducir los riesgos de la sobresegmentación e incrementar la robustez del algoritmo.

Referencias

Cuevas Jimenez, E. V. y Zaldivar Navarro, D. "Visión por computador utilizando Matlab y el Tollbox de procesamiento digital de imágenes". http://www.academia.edu/7448239/Visi%C3%B3n_por_Computador_utilizando_MatLAB_Y_el_Toolbox_de_Procesamiento_Digital_de_Im%C3%A1genes.

Fiter, E. L. (2012) "Descripción, comparación y ejemplos de uso de las funciones de la toolbox de procesado digital de imágenes de MATLAB". Proyecto fin de carrera. Escuela Universitaria de Ingeniería Técnica de Telecomunicaciones, Universidad Politécnica de Madrid.

Gonzales, M. A. y Ballerin, V. L. (2008). "Segmentación de imágenes utilizando la transformada de Watershed: obtención de marcadores mediante lógica difusa". IEEE Latin American Transaction, Vol. 6, N° 2. p 223 - 228.

La Serna Palomino, N. y Garcia Hilares, N. (2011) "Implementación del algoritmo de Watershed para el análisis de imágenes médicas". Revista de Investigación de Sistemas e Informática. Facultad de Ingeniería de Sistemas e Informática, Universidad Nacional Mayor de San Marcos. RISI 8(2), p 67 - 74. ISSN 1816-3823.

Mathworks. (1994-2018) "Image Processing Toolbox". https://es.mathworks.com/products/image.html ?s_tid=srchtitle

Rodriguez Medina, C. G. y Navas, G. S. (2014) "Aplicación del filtro de Canny en la esteganografía digital". XVI Workshop de Investigadores en Ciencias de la Computación - WICC 2014 Ushuaia. RedUNCI. p 806 - 811. ISBN 978-950-34-1084-4.